Real Time Stokes Inversion Using Multiple Support Vector Regression*

David Rees[1,2], Ying Guo[1], Arturo López Ariste[2,3], and Jonathan Graham[2]

[1] Centre for Intelligent Systems Design
CSIRO ICT Centre, PO Box 76, Epping, NSW 1710, Australia
Email: {david.e.rees, ying.guo}@csiro.au
[2] High Altitude Observatory, NCAR** PO Box 3000, Boulder, CO-80305, USA
[3] THEMIS-CNRS UPS 853 C/Via Lactea sn. 38200 La Laguna, Canary Islands, Spain

Abstract. Solution of the inverse problem to estimate the vector magnetic field distribution on the sun from the profiles of the Stokes parameters of polarised light in magnetically sensitive spectral lines is a vital task in our understanding of solar activity. Recently machine learning techniques such as principal component analysis and neural networks have led to the development of real time inversion computer codes. This paper is the latest contribution from the Australian-American-French Connection, an international team playing an active role in this Stokes inversion revolution. A new inversion method called Multiple Support Vector Regression is described and applied here for the first time to synthetic Stokes profile data.

1 Introduction

The magnetic field that permeates the external layers of the sun plays a fundamental role in solar activity. Estimation of the magnetic field distribution near the solar surface is done indirectly using spectropolarimetry, i.e. measurement of the wavelength dependence of the Stokes parameters (or Stokes profiles) of polarised radiation in magnetically sensitive spectral lines. The solution of the inverse problem to infer the vector magnetic field from Stokes profile data is known as Stokes inversion (Socas-Navarro [1]). Modern spectropolarimeters provide accurate measurements of the Stokes profiles of many different spectral lines formed at various atmospheric heights. Future space- and ground-based instruments, such as Solar-B and SOLIS, will achieve unprecedented spatial resolution and coverage of the solar surface. The expected flood of data from such instruments has recently been the catalyst for the development of several new approaches to Stokes inversion based on machine learning, aimed at real time data analysis.

** The National Center for Atmospheric Research is sponsored by the National Science Foundation

* Submitted January 30, 2004 to KES-2004. Send correspondence to David Rees.

Initial steps towards real time inversion were made by Rees et al [2][3] who proposed a database search method using principal component analysis (PCA). Socas-Navarro et al [4] and López Ariste et al [5] showed that so-called PCA inversion is over two orders of magnitude faster than traditional nonlinear least squares model fitting (Auer et al [6]). Essentially real time inversion has been achieved using neural networks (Carroll and Staude [7]; López Ariste et al [5]; Socas-Navarro [8]). In parallel with this PCA and neural network research we have been investigating an alternative approach which we call multiple support vector regression (MSVR) (Rees and Guo [9]). This is the focus of the current paper.

The rest of the paper is structured as follows. Section 2 summarises the MSVR method and Section 3 illustrates its application to synthetic unpolarised intensity profile data. Section 4 presents the first application of MSVR to synthetic Stokes profile data. We conclude in Section 5, setting the research agenda for the next stage of development of MSVR.

2 Multiple Support Vector Regression

In many applications support vector machines (SVMs) (Vapnik [10]) have been shown to outperform neural networks when applied to classification problems. SVMs have also been used for nonlinear regression. This is known as support vector regression (SVR). Cristianini and Shawe-Taylor [11] provide an excellent introduction to both SVMs and SVR. In this section we show how SVR can be used for parameter estimation. Since in general there are multiple parameters involved, we refer to this new method of inversion as multiple SVR or MSVR (Rees and Guo [9]).

In brief the inversion problem can be formulated as follows. Suppose we measure an N dimensional signal $\mathbf{S} = (S_1, \ldots, S_N)$ and associate with this signal a set of parameters $\mathbf{p} = (p_1, p_2, \ldots)$. Thinking of $\mathbf{S}$ as an operator (generally nonlinear) on $\mathbf{p}$, the goal is to find the inverse operator $\mathcal{F}$ such that

$$\mathbf{p} = \mathcal{F}(\mathbf{S}(\mathbf{p})) \tag{1}$$

To approximate $\mathcal{F}$ we use a *training set* of signals $\mathbf{S}_j = \mathbf{S}(\mathbf{p}_j), j = 1, \ldots, M$ corresponding to M different parameter sets $\mathbf{p}_j$. In many cases N is large and it is advantageous to reduce dimensionality by PCA, reconstructing $\mathbf{S}$ using eigenvectors estimated from this training set(Rees et al [2][3]). Then instead of $\mathbf{S}$ we can work with the vector $\mathbf{E} = (e_1, \ldots, e_n)$ of eigenfeatures or principal components, where $n \ll N$, and the inversion problem can be recast as finding $\mathcal{F}$ such that

$$\mathbf{p} = \mathcal{F}(\mathbf{E}(\mathbf{p})) \tag{2}$$

For a model with L physical parameters each eigenfeature vector $\mathbf{E}_i$ in the training set has an associated parameter set $\mathbf{p}_i = (p_{i1}, \cdots, p_{ik}, \cdots, p_{iL})$. In order to estimate a particular parameter p_k, we organise the M training examples as $(\mathbf{E}_1, p_{1k}), \cdots, (\mathbf{E}_i, p_{ik}), \cdots, (\mathbf{E}_M, p_{Mk})$, regarding $\mathbf{E}_i$ as input vectors and p_{ik} as the associated output values for application of the SVR algorithm.

The goal of SVR is to find a function $f_k(\mathbf{E})$ such that

$$|f_k(\mathbf{E}_j) - p_{jk}| \leq \epsilon, \text{ for } j = 1, \ldots, M, \tag{3}$$

where $\epsilon > 0$. Thus the function value $f_k(\mathbf{E}_j)$ has at most ϵ deviation from the actually obtained targets p_{jk} for all the training data, and at the same time, is as smooth as possible. The SVR function has the form:

$$f_k(\mathbf{E}) = \sum_{i=1}^{M} \alpha_{ik} K_k(\mathbf{E}_i, \mathbf{E}) + b_k. \tag{4}$$

where α_{ik} and b_k are constants, and K_k is the *kernel* function. The index k emphasises that one is free to choose different kernel functions for different system parameters.

For some cases linear SVR may be sufficient, but in general nonlinear SVR is desired. In the latter case a number of kernel functions have been found to provide good performance, including polynomials, radial basis functions (RBF), and sigmoid functions. The SVR optimisation problem is then solved in accordance with standard techniques (see, for example, Cristianini and Shawe-Taylor [11]).

The regression functions $p_k = f_k(\mathbf{E})$, for $k = 1, \ldots, L$, learned by this process, constitute the inverse operator $\mathcal{F}$ in equation (2).

3 Application to Unpolarised Spectra

We now illustrate the method using synthetic unpolarised intensity profiles modelled analytically by

$$I = 1 + \frac{1}{1 + \eta_0 e^{-(x/\delta)^2}} \tag{5}$$

where x is a dimensionless wavelength measured from line centre. This model has two adjustable parameters: η_0 which is the line to continuum opacity ratio at line centre and δ which controls the line broadening. In terms of the previous notation, $\mathbf{p} = (\eta_0, \delta)$ and $\mathbf{S}$ is composed of values of I sampled at N values of the wavelength x. The goal is to find the regression functions $\eta_0 = f_{\eta_0}(\mathbf{E})$ and $\delta = f_\delta(\mathbf{E})$.

3.1 One Parameter Case

Fixing $\delta = 1$, we generated a training set of $M = 19$ profiles using the opacity values $\eta_0 = 1 : (0.5) : 10$, i.e. from 1 to 10 in steps of 0.5. The profiles were computed at $N = 61$ wavelengths $x = -3 : (0.1) : 3$ and just two eigenfeatures, i.e. a 2–dimensional eigenfeature vector $\mathbf{E} = (e_1, e_2)$ was used. The training data were fitted with a polynomial kernel. One result of the fitting is automatic selection of the number of support vectors required for the SVR function. In this case there are 7 support vectors. The SVR function is a smooth interpolating function which can be used for accurate parameter estimation for any eigenfeature vector, not just those in the training set. The training set and the SVR function (a smooth interpolating curve) are shown in Figure 1.

Synthetic test data were generated for a large number of values of η_0. These test data were used as "observations" and the parameter values estimated with the SVR function. The errors in these estimated η_0 were found to be less than 1% for all test data.

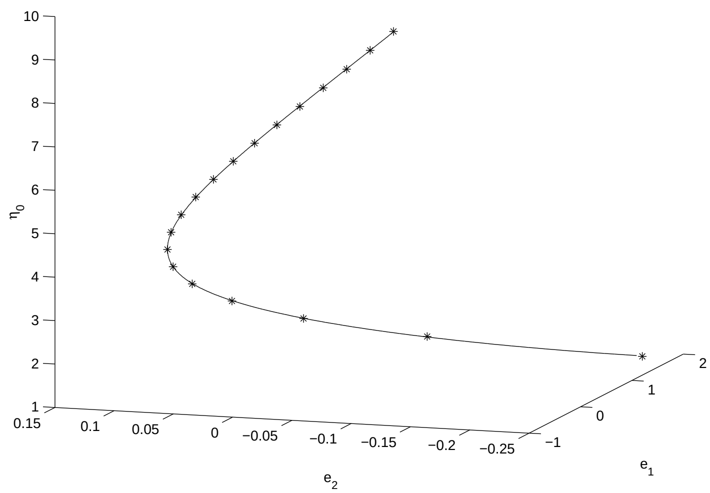

Fig. 1. Unpolarised training data (*) and SVR function $\eta_0 = f_{\eta_0}(\mathbf{E})$ (continuous curve) for case of fixed $\delta = 1$.

3.2 Two Parameter Case

Here we allow both parameters to vary, generating a training set of $M = 121$ profiles for $\eta_0 = 1 : (0.9) : 10$ and $\delta = 0.5 : (0.1) : 1.5$ for $N = 81$ wavelengths $x = -4 : (0.1) : 4$. We used a 3–dimensional eigenfeature vector $\mathbf{E} = (e_1, e_2, e_3)$ and fitted the regression functions with an RBF kernel.

The number of support vectors defining the regression functions $f_{\eta_0}(\mathbf{E})$ and $\delta = f_\delta(\mathbf{E})$ were 83 and 72 respectively. The training data and the regression functions (smooth interpolating surfaces) viewed as functions of e_1 and e_2 are shown in Figure 2.

Synthetic test data were again generated for a large number of parameter values and the regression functions were used to estimate the parameters from these "observations". The errors in these estimates were found to be less than 1.3% for η_0 and less than 0.3% for δ for all test data.

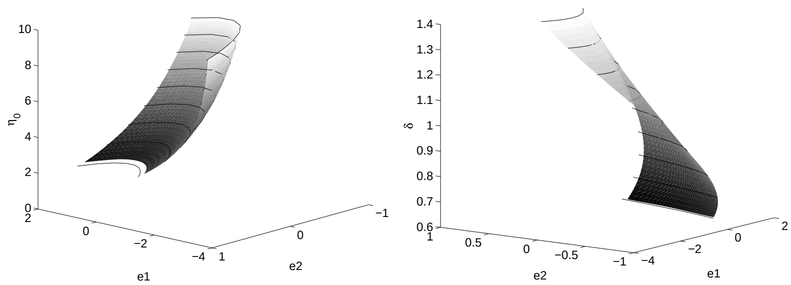

Fig. 2. Unpolarised SVRs (smooth interpolating surfaces) $\eta_0 = f_{\eta_0}(\mathbf{E})$ (*left*) and $\delta = f_\delta(\mathbf{E})$ (*right*). The level curves are defined by the training data

4 Application to Polarised Spectra

We now apply MSVR to invert Stokes profiles. For simplicity we consider only the spectral profiles of intensity I and net circular polarisation V. A training set of $M = 399$ synthetic I and V profiles sampled at 100 wavelengths was generated for a magnetically sensitive spectral line of neutral iron by solving the equations of polarised radiative transfer in a model of the solar photosphere for a range of magnetic field strengths $B = 0 : (100) : 2000$ G (Gauss), and inclinations $\gamma = 0 : (5) : 90$ degrees to the line of sight; the field azimuth was not varied. Thus in this model the parameter vector is $\mathbf{p} = (B, \gamma)$ and the signal vector is a 200-dimensional concatenation of the I and V profiles, which, on applying PCA separately to I and V and retaining only the first two eigenfeatures for each, leads to a composite 4-dimensional eigenfeature vector, $\mathbf{E} = (e_1, e_2, e_3, e_4)$.

The goal is to estimate the SVR functions $B = f_B(\mathbf{E})$ and $\gamma = f_\gamma(\mathbf{E})$. We fitted the regression functions with an RBF kernel. The number of support vectors varied depending on selection of certain fitting criteria in the SVR algorithm, but averaged about 80. Here we present only the results for $f_B(\mathbf{E})$. Training data and regression function (smooth interpolating surface) are shown in Figure 3 as functions of e_1 and e_2.

As in the unpolarised case synthetic test data were generated and used as "observations". The absolute errors in the estimates of B from these data, viewed as an error surface in Figure 4, were less than 10G, well under the errors typically found in analysis of real observational data.

5 Conclusion

To our knowledge MSVR inversion, as formulated in this paper, is completely novel. It is a generic technique which should be widely applicable to inverse problems in science, medicine, and engineering. Given that the method provides explicit functional representations of model parameters, as does nonlinear regression by a neural network, MSVR is a new option for real time data processing.

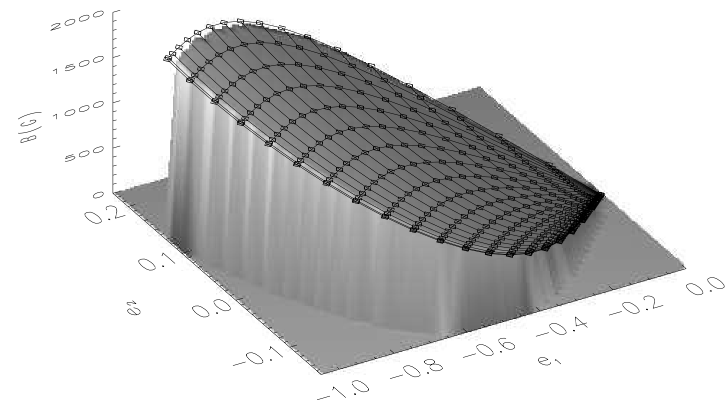

Fig. 3. Polarised training data viewed as a dark mesh superposed on SVR function $B = f_B(\mathbf{E})$ (smooth interpolating surface) for magnetic field strength.

Our very preliminary tests indicate that MSVR will indeed work for Stokes inversion, but much more detailed research and testing are required before MSVR could be said to be a viable alternative to neural network inversion. Issues to be addressed include how best to form the composite signal and associated eigenfeature vectors, especially when all four Stokes profiles are involved, i.e. linear as well as circular polarisation are treated simultaneously.

Neural network inversion is currently emerging as the method of choice for on-board real time data processing, for example on the Helioseismic Magnetic Imager (HMI) experiment on the Solar Dynamics Observatory mission to be launched in 2007. It is worth noting here that HMI is a filtergraph instrument and samples the Stokes spectra at only a small number of wavelengths. Graham et al [12] showed that even with such limited wavelength coverage it is possible to obtain reliable vector magnetic field estimates by traditional inversion using nonlinear least squares model fitting. Obviously in this case the signal data already is low dimensional and the PCA compression step discussed in this paper is not necessary. It will be interesting to investigate the application of MSVR to such data.

Acknowledgements

David Rees gratefully acknowledges the generous support of the Affiliate Scientist Program at HAO.

References

1. Socas-Navarro H.: Stokes inversion techniques: Recent achievements and future horizons, in: Advanced Solar Polarimetry – Theory, Observation and Instrumentation, ed. M. Sigwarth, (2001)

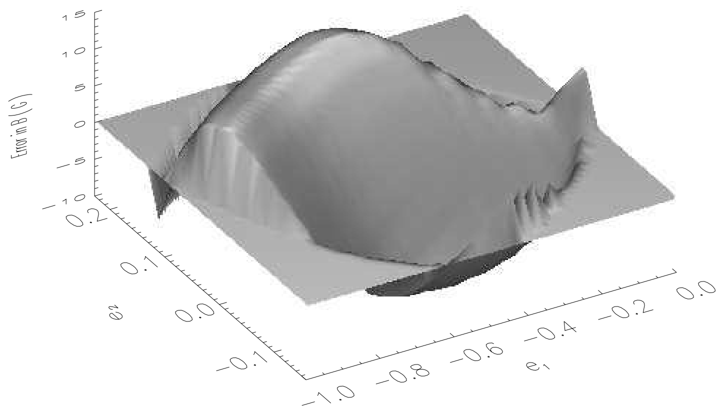

Fig. 4. Error surface for field strengths estimated with the SVR function $B = f_B(\mathbf{E})$

2. Rees, D. E., Mulcahy, D., and Thatcher, J. : A database technique for fast spectral line inversion, Proc. IASTED Int. Conference on Signal Processing and Communications, Canary Islands (1998) 408–411
3. Rees D., López Ariste A., Thatcher J., Semel M.: Fast inversion of spectral lines using principal component analysis. I. Fundamentals, A&A **355** (2000) 759–768.
4. Socas-Navarro H., López Ariste, A., Lites B. W.: Fast inversion of spectral lines using principal component analysis. II. Inversion of real Stokes data, ApJ **553** (2001) 949-954.
5. López Ariste A., Rees, D., Socas-Navarro, H., Lites, B.W.: Pattern recognition techniques and the measurement of solar magnetic fields, Proc.SPIE Vol 4477 Astronomical Data Analysis, San Diego, August (2001) 96–106.
6. Auer, L. H., Heasley, J. M., House, L. L.: The determination of vector magnetic fields from Stokes profiles. Solar Phys. **55** (1977) 47–61
7. Carroll, T., and Staude, J. : The inversion of Stokes profiles with artificial neural networks. A&A **378** (2001) 316 – 326
8. Socas-Navarro, H.: Measuring solar magnetic fields with artificial neural networks. Neural Networks **16** (2003) 355–363
9. Rees, D.E., and Guo, Y.: Ghosts in the machine: from turbulence, faces and ducks to solar magnetic fields, in: Trujillo Bueno J., Sanchez Almeida J. (eds.), Proceedings of the Solar Polarisation Workshop 3, Tenerife, Oct 2002, ASP Conf. Series (2003)
10. Vapnik, V. N.: The Nature of Statistical Learning Theory. Springer, NY (1995)
11. Cristianini, N., and Shawe-Taylor: An Introduction to Support Vector Machines, Cambridge University Press, Cambridge, UK (2000)
12. Graham, J.D, López Ariste, A., Socas-Navarro H., Tomczyk, T.: Inference of solar magnetic field parameters from data with limited wavelength sampling. Solar Phys. **208** (2002) 211–232